JOURNAL OF Evolutionary Biology

doi: 10.1111/j.1420-9101.2012.02537.x

SHORT COMMUNICATION

Contrasting patterns of selection acting on MHC class I and class II DRB genes in the Alpine marmot (*Marmota marmota*)

K. KUDUK*, A. JOHANET†, D. ALLAINÉ†, A. COHAS† & J. RADWAN*

*Institute of Environmental Sciences, Jagiellonian University, Kraków, Poland
†Université de Lyon, Université Lyon 1, CNRS UMR 5558, laboratoire de Biométrie et Biologie Evolutive, Villeurbanne, France

Keywords:

bottleneck;
duplication;
parasites;
polymorphism;
positive selection;
rodents;
Sciuridae.

Abstract

The major histocompatibility complex (MHC) genes code for proteins that play a critical role in the immune system response. The MHC genes are among the most polymorphic genes in vertebrates, presumably due to balancing selection. The two MHC classes appear to differ in the rate of evolution, but the reasons for this variation are not well understood. Here, we investigate the level of polymorphism and the evolution of sequences that code for the peptide-binding regions of MHC class I and class II DRB genes in the Alpine marmot (*Marmota marmota*). We found evidence for four expressed MHC class I loci and two expressed MHC class II loci. MHC genes in marmots were characterized by low polymorphism, as one to eight alleles per putative locus were detected in 38 individuals from three French Alps populations. The generally limited degree of polymorphism, which was more pronounced in class I genes, is likely due to bottleneck the populations undergone. Additionally, gene duplication within each class might have compensated for the loss of polymorphism at particular loci. The two gene classes showed different patterns of evolution. The most polymorphic of the putative loci, *Mama-DRB1,* showed clear evidence of historical positive selection for amino acid replacements. However, no signal of positive selection was evident in the MHC class I genes. These contrasting patterns of sequence evolution may reflect differences in selection pressures acting on class I and class II genes.

Introduction

The major histocompatibility complex (MHC) plays a crucial role in the adaptive immune response of vertebrates. The MHC is a multigene family encompassing classical class I and class II genes, as well as nonclassical MHC genes and pseudogenes (Kelley *et al.*, 2005). MHC class I and class II comprise several loci coding for proteins involved in antigen presentation and initiation of the T-cell-mediated specific immune responses. Antigens are presented in a groove formed by proteins coded by second and third exons of MHC class I genes, whereas in MHC class II the groove is a dimer formed by products of second exons of two separate genes (DRA and DRB, DQA and DQB, etc.). Class I molecules are involved mainly in the immune response to intracellular pathogens, such as viruses, whereas class II molecules are involved in the recognition of extracellular pathogens (Janeway *et al.*, 2004), although there is some overlap of these functions (Iwasaki & Medzhitov, 2010).

The MHC genes are the most polymorphic genes known in vertebrates, with the majority of polymorphism located at residues coding for the antigen binding sites (ABS, Garrigan & Hedrick, 2003; Janeway *et al.*, 2004). At these sites, an excess amount of nonsynonymous substitutions ($d_N/d_S > 1$), which is a hallmark of positive selection for amino acid substitutions, is observed (Bernatchez & Landry, 2003; Garrigan & Hedrick, 2003; Piertney & Oliver, 2006). Alleles in different species can be more similar to each other than some

Correspondence: Jacek Radwan, Institute of Environmental Sciences, Jagiellonian University, Gronostajowa 7, 30-387 Kraków, Poland.
Tel.: +48 12 6646849; fax: +48 126646912;
e-mail: j.w.radwan@gmail.com

alleles within species (trans-species polymorphism), indicating long-term maintenance of allelic lineages (Klein *et al.*, 1998). The high number of alleles, their high divergence, the molecular signatures of positive selection and trans-specific polymorphism imply the maintenance of genetic variability via balancing selection (Takahata & Nei, 1990). The rapid evolution of parasites may contribute to this selection as the parasites' adaptation to common host genotypes may result in a negative frequency-dependent selection on the MHC alleles (Snell, 1968). Additionally, different MHC alleles may be favoured at different times depending on the composition of the parasite community (Hedrick, 2002). Furthermore, heterozygous individuals may gain an advantage in viability by being able to present a wider range of antigens (Doherty & Zinkernagel, 1975). Finally, MHC variation can also be maintained by mate choice (Hedrick, 1992).

While the above mechanisms are invoked in explaining polymorphism of both MHC classes, class I genes appear to undergo a faster rate of birth-and-death than class II loci among mammals (Hughes & Nei, 1989; Yeager & Hughes, 1999). In contrast, in salmonids and in the Lake Tana large barb species flock, MHC class I lineages are more divergent and older than class II genes (Kruiswijk *et al.*, 2005, Shum *et al.*, 2001). The reasons for this difference between class I and class II genes are not well understood, but are likely to result from differences in selection pressures acting on both MHC classes. Such differences should be reflected in the rates of sequence evolution, but such comparisons have rarely been made between MHC classes. In mice, d_N/d_S in MHC class I genes is considerably lower than that in MHC class II (Hughes & Nei, 1988, 1989), consistent with stronger balancing selection maintaining class II lineages for longer evolutionary time. Selection differences were also found for *Labeobarbus intermedius* species flock, with the lower number of sites showing elevated d_N/d_S values in class I relative to class II (Wegner, 2008). In humans, however, differences in the rate of the birth-and-death process are not mirrored in the differences in the strength of positive selection, as d_N/d_S ratios are very similar for both MHC classes (Hughes & Nei, 1988, 1989).

Here, we characterize, for the first time, the variation in the sequences of second exon of MHC class I gene and MHC class II DRB gene in the Alpine marmot (*Marmota marmota*), and compare the extent of historical selection acting on the two MHC classes. The high parasite load found in Alpine populations of this species (Callait, 1999) suggests that immune resistance could be a crucial factor in determining individual fitness and that it may have consequences for population viability. Furthermore, MHC is likely to play a role in mate choice in this species. The Alpine marmot is a socially monogamous species with a high level of extra-pair paternity (Cohas *et al.*, 2006), and previous studies showed significant influence of individuals' genetic characteristics on extra-pair mate choice (Cohas *et al.*, 2006, 2007, 2008).

Material and methods

We obtained hair samples from 38 individuals from three populations (La Grande Sassière Nature Reserve – $n = 15$, the National Park of Les Ecrins – $n = 10$ and the Maurienne valley – $n = 13$, described in Goossens *et al.*, 2001). Genomic DNA was extracted from 15 to 30 hairs by incubation at 66 °C for 80 min in 50 μL lysis buffer (20 mM Tris–HCl, 1.5 mM $MgCl_2$, 25 mM KCl, 0.5% Tween 20 and 0.1 mg mL^{-1} proteinase K), followed by 20 min of proteinase K inactivation at 96 °C (Cohas *et al.*, 2008). DNA amplifications were performed in the Mastercycler (Eppendorf, Hamburg, Germany) in 10 μL reaction mixture containing 5 μL of HotStarTaq Polymerase Master Mix (Qiagen, Hilden, Germany), 0.2 μL of both primers at 100 μM, 3.6 μL of water and 1 μL of DNA at a concentration of 30 ng μL^{-1}. The cycling scheme was 95 °C for 15 min, followed by 35 cycles at 95 °C for 30 s, primer-specific annealing temperature (Table S2) for 30 s, 72 °C for 60 s and a final extension step at 72 °C for 10 min.

To investigate the expression of the MHC loci in marmots, we also extracted RNA from RNAlater (Qiagen)-preserved spleens of three individuals coming from the Maurienne valley using the RNeasy Mini Kit (Qiagen) and reverse-transcribed the samples using Omniscript RT Kit (Qiagen) following the manufacturer's protocols.

MHC class I primers were designed in conserved regions of the second exon based on rodent sequences obtained from GenBank (Table S1). PCR products obtained with the primers from a subsample of 18 individuals were then purified using the MinElute PCR Purification Kit (Qiagen) and cloned using StrataClone PCR Cloning Kit (Agilent Technologies, Santa Clara, CA, USA). Twenty-four clones per pool were amplified using the M13 forward and reverse primers and sequenced with BigDye Terminator 3.1 sequencing kit (Applied Biosystems, Carlsbad, CA, USA), and products were sequenced on an ABI 3130xl Genetic Analyzer (Applied Biosystems). Phylogenetic analysis grouped the obtained sequences (199 or 202 bp) into four strongly divergent branches, allowing design of cluster-specific primers (MarmR1, MarmR2, MarmR3 and MarmR4) amplifying 175- to 180-bp-long sequences (Table S2). Primer sequences, annealing temperature and concentration for each primer set are specified in Tables S1 and S2.

For initial amplification of the second exon of MHC class II DRB loci, we used conservative primers located in the first (EV2F) and the third (EV1R, Table S1) exons, designed using the GenBank mammalian sequences (Kloch, unpublished data). Amplification from gDNA was not successful, and therefore, we amplified each cDNA sample. PCR products were cloned and sequenced as previously described to obtain 16 sequences from each cDNA sample. The sequences formed two clades (Fig. 2), and consequently, two sets of marmot-specific primers were designed: forward MM_DRB_F1 and reverse

MM_DRB_R3, and forward MM_DRB_F2 and reverse MM_DRB_R2, amplifying a 203- to 204-bp fragment of the second exon of the DRB locus (Table S1 and S2).

We used several approaches for genotyping individuals and determining sequences of individual alleles. For genotyping, we performed the single strand conformational polymorphism (SSCP) using capillary electrophoresis (as described in Biedrzycka & Radwan, 2008). The homozygotes were sequenced directly, and sequences of heterozygotes with novel SSCP patterns were inferred based upon previous knowledge of migration patterns of known sequences (from homozygotes and previously sequenced clones). Sequences of all such obtained alleles were confirmed by 454 sequencing (Babik *et al.*, 2009; Zagalska-Neubauer *et al.*, 2010) of other sample of individuals (A. Johanet, A. Cohas & D. Allainé, unpublished data). The nomenclature principles proposed by Klein *et al.* (1990) were followed to designate allele names for both MHC classes (see Table S2).

The phylogenetic trees were constructed under the Bayesian approach with MRBAYES3.1.2 (Ronquist & Huelsenbeck, 2003). For the computation of both trees, the likelihood settings corresponded to the general time-reversible (GTR) model + Γ with the parameter values estimated from the data. Priors were set to default values. Four Metropolis-coupled Markov-chain Monte Carlo simulations (three of them 'heated', temperature = 0.20) were run for 5×10^6 generations and sampled every 1000 generations. The first 1000 trees were discarded as a burn-in. To calculate the posterior probability of each bipartition, the majority-rule consensus tree was computed from the 4000 sampled trees. Pseudogene sequence was excluded from tree construction.

Observed (Ho) and expected (He) heterozygosity, Hardy–Weinberg equilibrium tests and allele frequencies were computed for each locus separately using the ARLEQUIN 3.5 software (Excoffier & Lischer, 2010). Allelic richness (R) for each locus was calculated using the rarefaction method implemented in FSTAT 2.9.3 (Goudet, 2001).

The average pairwise nucleotide distances (Kimura 2-parameter model, K2P), the Poisson-corrected amino acid distances and average rates of synonymous (d_S) and nonsynonymous (d_N) substitutions per site, using the Nei–Gojobori method (Nei & Gojobori, 1986) with the Jukes–Cantor correction for multiple substitutions, were computed in MEGA5 software (Tamura *et al.*, 2011). The locations of putative ABS sites in MHC class I and MHC class II DRB genes were inferred from the structure of human HLA genes (Reche & Reinherz, 2003).

We tested for positive selection shaping the evolution of the second exon of MHC class I and MHC class II DRB genes using two approaches. First, we compared d_S and d_N for all sites, ABS and non-ABS, using the *Z*-test in MEGA5 software. The standard error of d_N and d_S was obtained through 1000 bootstrap replicates (Nei & Kumar, 2000). Second, we used CODEML, included in PAML 4.2 (Yang, 2007), to compare the likelihoods of three codon-based models of sequence evolution. These models are the following: M0, a one ω (i.e. d_N/d_S) ratio; M7, a nearly neutral ratio ($0 < \omega < 1$) with ω-ratio variation depending on β-distribution; and M8, a positive selection ratio (some sites evolve with $\omega > 0$) with ω-ratio variation depending on β-distribution. We chose the best-fitting model on the basis of the lowest value of the Akaike Information Criterion (AIC, Posada & Buckley, 2004). Positively selected codons under the M8 model were identified through the Bayes empirical Bayes procedure (Zhang *et al.*, 2005).

Results and discussion

In total, we found 7 MHC class I sequences and 11 MHC class II DRB sequences. The sequences were deposited in GenBank (MHC class II; accession numbers JQ837902–JQ837912) and provided in Supporting Information (MHC class I). All except one sequence (*Mama-DRB1**08) lacked stop codons or indels. Bayesian phylogenetic trees constructed for the second exon of MHC class I genes and MHC class II DRB genes of rodent species are shown in Figs 1 and 2. The phylogenetic analysis of the MHC class I alleles distinguished two well-supported clusters of marmot sequences and two additional highly divergent sequences. The sequences of marmot MHC class II DRB genes formed two highly supported and divergent clusters. The high divergence of both the MHC class I and MHC class II clusters suggests that they may constitute separate loci. Per individual, four to six unique nucleotide sequences were found in MHC class I genes, corresponding to four to five amino acid sequences, and two to four nucleotide sequences were found in MHC class II DRB genes, corresponding to two to three amino acid sequences. Analysis of the genotypes of all individuals supported the notion that each cluster (amplified by a specific primer pair) represents a separate locus. Firstly, none of the cluster-specific primer pairs amplified more than two alleles per individual. Secondly, at least one sequence belonging to each cluster was present in all individuals. Amplification and sequencing of cDNA samples confirmed the expression of alleles belonging to all of the clusters detected in gDNA, for both the MHC class I and MHC class II DRB genes (Figs 1 and 2). Allele *Mama-DRB1**08 contained an insertion altering the reading frame. After removing the insertion, sequence aligned fully with *Mama-DRB1**02 allele, and it never occurred along with more than one other DRB1 sequence, so it most likely belonged to the same locus. It has not occurred in individuals sampled for MHC expression, so it is unclear whether it is transcribed. Our data thus indicate the presence of four MHC class I loci (two represented by a single allele) and two MHC class II DRB loci in each of the investigated individuals.

MHC class I clusters did not show well-supported relationships to other rodent sequences (Fig. 1), consistent

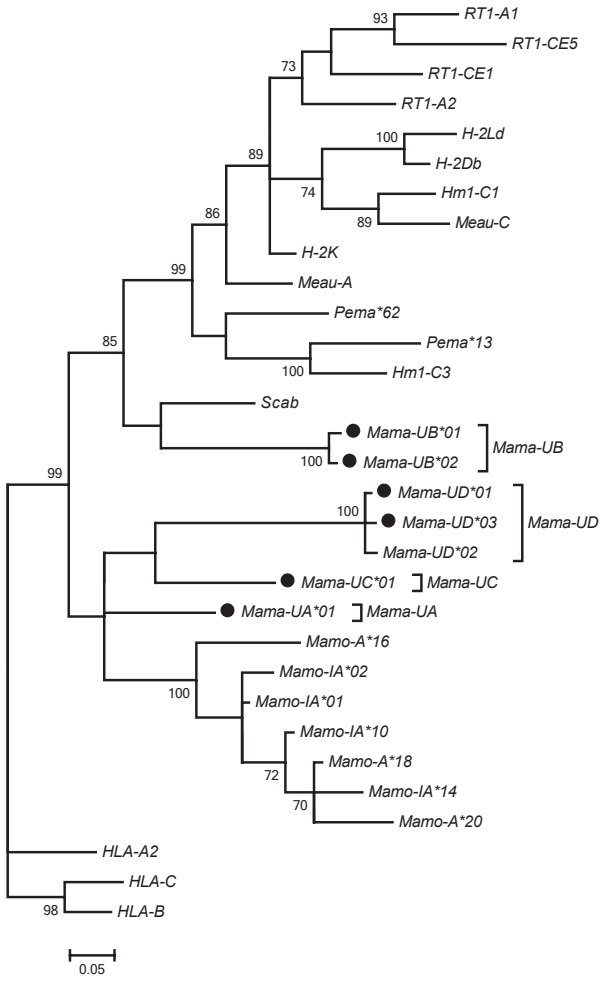

Fig. 1 A phylogenetic tree constructed from a 174-bp fragment of the second exon of MHC class I gene illustrating the localization of the *Marmota marmota* alleles (*Mama-U*) as compared to other rodent alleles. Bayesian posterior probabilities above 70% are shown next to branches. The rodent sequences used to generate the tree are as follows: *Rattus norvegicus*: *RT1-A1* (NM_001008827.1), *RT1-A2* (BC088465.1), *RT1-CE1* (NM_001008832.2), *RT1-CE5* (BC098841.1); *Mus musculus musculus*: *H-2D-b* (M18523.1), *H-2K* (V00746.1), *H-2Ld* (V00752.1); *Cricetulus griseus*: *Hm1-C1* (AY064386.1), *Hm1-C3* (AY064388.1); *Mesocricetus auratus*: *Meau-A* (M63144.1), *Meau-C* (M63147.1); *Peromyscus maniculatus*: *Pema*62* (U12887.1), *Pema*13* (U12822.3); *Sciurus aberti aberti*: *Scab* (M97617.1); *Marmota monax*: *Mamo-IA*01* (AF153907.1), *Mamo-IA*02* (AF146093.1), *Mamo-IA*10* (AF146091.1), *Mamo-IA*14* (AF154348.1), *Mamo-A*16* (AY139187.1), *Mamo-A*18* (AY139189.1), *Mamo-A*20* (AY139191.1). *Homo sapiens* sequences: *HLA-A2* (K02883.1), *HLA-B* (AJ292075.2) and *HLA-C* (Y18533.1) were used to root the tree. The GenBank accession numbers are in parentheses. Filled circles denote *M. marmota* expressed alleles.

with fast gene turnover characteristic of this class in mammals (Hughes & Nei, 1988, 1989; Yeager & Hughes, 1999). Within MHC class II, *Mama-DRB1* cluster grouped with sequences of *Spermophilus suslicus* DRB locus and *Mama-DRB2* cluster grouped with a *Dipodomys spectabilis* DRB sequence (Fig. 2). Although relatively short sequence analysed does not allow drawing firm conclusions about time-order of duplications, especially that strong positive selection may interfere with phylogenetic inference (Wettstein *et al.*, 1996), the grouping of DRB2 with *Dipodomys* suggests that duplication occurred before the split of Sciuromorpha from other rodent suborders and that Mama-*DRB2* ortholog has subsequently been lost in *Spermophilus*.

Our phylogenetic tree contained other *Marmota* (MHC class I) or *Spermophilus* (MHC class II) sequences, but despite relatively short time since the split between marmot and ground-squirrel lineages (some 10 Mya; Armitage, 2000), we observed no indications of trans-species polymorphism: the alleles within clusters were more similar to each other than to alleles of other sciurid species (Figs 1 and 2). Whether trans-species polymorphism is observed between *Marmota* genera at MHC class II awaits data on other species.

Enhancing antigen-presenting capabilities by duplication of genetic loci is widespread among mammals (Kelley *et al.*, 2005). Among rodents, duplication of the MHC class I region has been reported in Muridae (Günther & Walter, 2000; Kumánovics *et al.*, 2002) and Cricetidae (Crew *et al.*, 1990; Lobigs *et al.*, 1995), where two to three loci were found, although our study is the first to document multiple MHC class I loci in Sciuridae (Wettstein *et al.*, 1996; Yang *et al.*, 2000). The MHC class II DRB region does not seem to be duplicated in most of the species studied thus far (e.g. mouse – *Mus musculus*, rat – *Rattus norvegicus*, Figueroa *et al.*, 1990; striped mouse – *Rhabdomys pumilio*, Froeschke & Sommer, 2005), including sciurids (spotted suslik – *S. suslicus*, Biedrzycka & Radwan, 2008; European ground squirrel – *Spermophilus citellus*, Ricanova *et al.*, 2011), and has been completely deleted in mole rats (Nizetić *et al.*, 1987). However, extensive duplications have also been reported (banner-tailed kangaroo rat – *D. spectabilis*, Busch *et al.*, 2008; bank vole – *Myodes glareolus*, Babik *et al.*, 2009). Unlike in some other mammals with duplicated loci (e.g. bank vole – *M. glareolus*, Babik *et al.*, 2009; macaque – *Macaca mulatta*, Khazand *et al.*, 1999), DRB in Alpine marmots does not show interindividual variation in the number of copies.

The frequencies of alleles in polymorphic putative loci are shown in Table S3. We found no deviations from Hardy–Weinberg equilibrium except for a significant excess of homozygotes in La Maurienne population at the *Mama-UD* locus (Table S4). This could result from segregation in this population of an allele which is not captured by our primers, by selection increasing frequency of homozygotes or by nonrandom sampling, for example resulting from family structure. These possibilities need to be resolved using larger sample sizes, although given that our primers were designed based on an extensive knowledge of variation in marmot

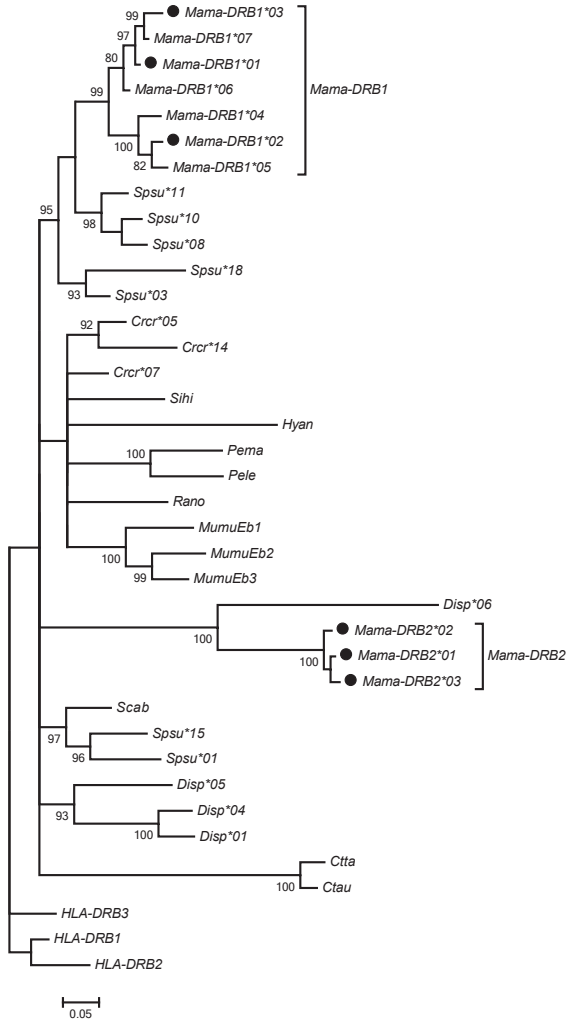

Fig. 2 A phylogenetic tree constructed from a 198-bp fragment of the second exon of MHC class II DRB gene illustrating the localization of the *Marmota marmota* alleles (*Mama-DRB*) among other rodent alleles. Bayesian posterior probabilities above 70% are shown next to branches. The rodent sequences used to generate the tree are as follows: *Spermophilus suslicus*: *Spsu*01* (EF569186.2), *Spsu*03* (JF681148.1), *Spsu*08* (EF569193.2), *Spsu*10* (EF569195.2), *Spsu*11* (EF569196.2), *Spsu*15* (EF569200.2), *Spsu*18* (HM461913.1); *Cricetus cricetus*: *Crcr*05* (AJ490315.1), *Crcr*07* (AJ490317.1), *Crcr*14* (AJ490324.1); *Sigmodon hispidus*: *Sihi* (AY169006.1); *Hypogeomys antimena*: *Hyan* (AJ416074.1); *Peromyscus maniculatus*: *Pema* (AF300855.1); *Peromyscus leucopus*: *Pele* (AF300851.1); *Rattus norvegicus*: *Rano* (AJ003226.1); *Mus musculus musculus*: *MumuEb1* (U88928.1), *MumuEb2* (U88929.1), *MumuEb3* (U88931.1); *Dipodomys spectabilis*: *Disp*01*(EU817479.1), *Disp*04* (EU817482.1), *Disp*05* (EU817483.1), *Disp*06* (EU817484.1); *Sciurus aberti aberti*: *Scab* (M97616.1); *Ctenomys talarum*: *Ctta* (DQ380442.1); *Ctenomys australis*: *Ctau* (GQ497466.1). The other sequences included are as follows: *Homo sapiens*: *HLA-DRB1* (AJ297583.1), *HLA-DRB2* (M17377.1) and *HLA-DRB3* (NM_022555.3). *HLA-DRB1*, *HLA-DRB2* and *HLA-DRB3* were used to root the tree. The GenBank accession numbers are in parentheses. Filled circles denote *M. marmota* expressed alleles.

sequences, nonamplification of some alleles does not appear likely. After translation to amino acid sequences, only one MHC class I cluster (*Mama-UD*, composed of two amino acid sequences) and one MHC class II DRB cluster (*Mama-DRB1*, composed of seven amino acid sequences) contained polymorphic protein sequences (Table S5).

The average pairwise nucleotide distances and amino acid distances were similar between MHC class I and class II sequences (Table 1). The MHC class I sequences were characterized by similar distances at ABS and non-ABS sites; however, for the MHC class II sequences, the distances were about four times higher at the ABS sites (Table 1). The two MHC classes apparently differ in the degree of historical selection undergone. No signal indicating positive selection was detected for MHC class I. The d_N/d_S ratio was less than one for both ABS and non-ABS sites (Table 1). The same result was obtained when the MHC class I cluster polymorphic in terms of amino acid sequences, the *Mama-UD* putative locus, was analysed separately (Table S6). Similarly, models of codon evolution for MHC class I sequences indicated that the nearly neutral model fitted the data the best (Table S7).

In contrast, we found support for historical selection favouring substitutions altering the amino acid sequences in the MHC class II DRB gene. When all sequences were considered, d_N significantly exceeded d_S in putative ABS sites but not in non-ABS sites (Table 1). While ABS positions are likely to differ to some extent among species (see below), this result suggests that positive selection was acting primarily on functionally important sites (Bernatchez & Landry, 2003). Model selection in PAML did not show a significantly better fit for the positive selection model (Table S7). However, this could be due to the high interlocus sequence divergence resulting in a saturation of both synonymous and nonsynonymous substitutions (Richman *et al.*, 2001). Indeed, when the *Mama-DRB1* cluster was analysed separately, the *Z*-test for positive selection was significant for ABS sites and also across all sites (Table S6). The M8 model incorporating positive selection provided the best fit for the data (Table S7). The Bayes empirical Bayes procedure identified six codons evolving under positive selection (posterior probability > 95%), three of which were located in the putative ABS (Table S5). This result suggests that functionally important sites do not exactly overlap between distantly related species, as also noted by other authors (e.g. Promerova *et al.*, 2009).

We thus found evidence for historical selection acting on the MHC class II genes but not on MHC class I genes. Weaker selection on MHC class I has also been reported for mice (Hughes & Nei, 1988, 1989), and it may be one of the factors facilitating faster turnover of this genes among mammals.

Diversity at both MHC class I and class II DRB genes is remarkably low compared to other rodents (reviewed in Goüy de Bellocq *et al.*, 2008). Such a low diversity in Alpine marmots can result from undergoing several

Table 1 The average pairwise nucleotide distances (Kimura 2-parameter models, K2P), Poisson-corrected amino acid distances (aa distance), the average rates of synonymous (d_S) and nonsynonymous (d_N) substitutions per site and the results of *Z*-test of positive selection for the MHC class I and MHC class II DRB in the second exon in *Marmota marmota*. Pairwise nucleotide distances, Poisson-corrected amino acid distances, d_S and d_N values are given as percentages per site. The standard error is in parentheses and was obtained through 1000 bootstrap replicates.

Sites	No. codons	K2P distance	aa distance	d_N	d_S	*Z*	*P*
MHC class I – all alleles							
All	58	20.6 (2.7)	34.4 (6.0)	17.5 (3.0)	31.4 (7.9)	−1.587	1.000
ABS	7	28.3 (11.1)	58.6 (22.9)	29.9 (13.5)	19 (16.8)	0.684	0.248
Non-ABS	51	19.8 (2.8)	31.8 (6.1)	16.1 (3.1)	34 (9.5)	−1.696	1.000
MHC class II DRB – all alleles							
All	67	22.5 (2.8)	32.9 (5.5)	20.1 (3.7)	31.3 (10.2)	−1.470	1.000
ABS	15	50.2 (13.0)	93.2 (21.9)	55.9 (17.0)	35.7 (18.5)	2.020	0.023
Non-ABS	52	17.0 (2.6)	22.8 (4.8)	13.6 (3.2)	30.3 (10.9)	−1.919	1.000

ABS, antigen binding sites.

bottleneck events (Preleuthner & Pinsker, 1993; Rassmann *et al.*, 1994). The low polymorphism observed at neutral microsatellite loci in Alpine marmots (Goossens *et al.*, 2001) confirms that bottleneck events are the most likely explanation for the low MHC variation in this species. While in historically bottlenecked populations considerable depletion in MHC variation is often observed (reviewed in Radwan *et al.*, 2010), strong balancing selection can modulate effects of drift (Aguilar *et al.*, 2004). Higher polymorphism levels were observed in the Alpine marmot's MHC class II DRB genes, compared to the MHC class I genes (eight alleles in the most polymorphic *Mama-DRB1* putative locus vs. three alleles in the *Mama-UD* putative locus). This is also consistent with stronger balancing selection acting on MHC class II DRB locus. Such selection should have resulted in both a higher prebottleneck diversity and retention of more DRB alleles during the bottleneck.

A large amount of evidence supports the maintenance of variation in MHC genes via pathogen-induced balancing selection (review in Spurgin & Richardson, 2010). However, different MHC genes may experience different selection pressure from parasites, as illustrated by divergent associations with parasite load between DQA and DRB genes in *Arvicola terrestris* (Tollenaere *et al.*, 2008). The weak selection acting on MHC class I genes in marmots could be caused by a weak intracellular pathogenic pressure. Unfortunately, little is known about pressure exerted by intracellular parasites in Alpine marmots, so this explanation needs to be assessed in future research. Additionally (or alternatively), the high number of expressed loci already present, with very divergent alleles, may be sufficient to recognize most of the intracellular pathogens. On the contrary, the positive selection acting on the DRB genes could result from higher extracellular pathogenic pressure acting on marmots. Indeed, autopsies revealed a huge burden of three common extracellular parasite species with a prevalence ranging from 32% to 89% and a total parasite biomass reaching, on average, 1% of the host body weight at the entry into hibernation in the population of La Sassière (Callait, 1999). Comparative evidence suggests that among rodents the load of extracellular parasite can indeed have strong effect on the evolution of MHC genes: Goüy de Bellocq *et al.* (2008) found that helminth species richness is a significant predictor of MHC class II polymorphism.

Alternatively, a higher level of selection on the MHC class II loci could result from an involvement in mate choice. The relative roles of MHC classes in producing olfactory cues are not well understood, although class I proteins have been shown to affect odours in mice (Penn & Potts, 1998; Leinders-Zufall *et al.*, 2004). Future work examining contemporary selection on MHC, resulting both from parasites and from mate choice, should provide some indications concerning the reasons for differences in modes of evolution in class I and class II DRB genes in Alpine marmots.

Acknowledgments

We warmly thank all students who were involved in catching marmots. We also extend thanks to the authorities of the Vanoise National Park for granting us permission to work in La Grande Sassière Nature Reserve. We thank Wiesław Babik for the fruitful discussions and advice. A post-doctoral fellowship was granted to A.J. by the University Claude Bernard Lyon 1. D.A. and A.C. were supported by the Agence Nationale de la Recherche (ANR, project ANR-08-BLAN-0214-03). J.R. and K.K. were supported by the Foundation for Polish Science (professor subsidy 9/2008) and Jagiellonian University DS/WBINOZ/INOS/762/10.

References

Aguilar, A., Roemer, G., Debenham, S., Binns, M., Garcelon, D. & Wayne, R. 2004. High MHC diversity maintained by balancing selection in an otherwise genetically monomorphic mammal. *Proc. Natl Acad. Sci. USA* **101**: 3490–3494.

Armitage, K.B. 2000. The evolution, ecology, and systematics of marmots. *Oecol. Mont.* **9**: 1–18.

Babik, W., Taberlet, P., Ejsmond, M.J. & Radwan, J. 2009. New generation sequencers as a tool for genotyping of highly polymorphic multilocus MHC system. *Mol. Ecol. Resour.* **9**: 713–719.

Bernatchez, L. & Landry, C. 2003. MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *J. Evol. Biol.* **16**: 363–377.

Biedrzycka, A. & Radwan, J. 2008. Population fragmentation and major histocompatibility complex variation in the spotted suslik, *Spermophilus suslicus*. *Mol. Ecol.* **17**: 4801–4811.

Busch, J.D., Waser, P.M. & DeWoody, J.A. 2008. Characterization of expressed class II MHC sequences in the banner-tailed kangaroo rat (*Dipodomys spectabilis*) reveals multiple DRB loci. *Immunogenetics* **60**: 677–688.

Callait, M.P. 1999. Relations hôte-parasite dans un écosystème alpin: dynamique annuelle et structures spatiales chez la marmotte alpine (*Marmota marmota*, Linné 1758). In *Université de Savoie*. Thèse de l'Université de Savoie, Chambéry, 281 pp.

Cohas, A., Yoccoz, N.G., Da Silva, A. *et al.* 2006. Extra-pair paternity in the monogamous alpine marmot (*Marmota marmota*): the roles of social setting and female mate choice. *Behav. Ecol. Sociobiol.* **59**: 597–605.

Cohas, A., Yoccoz, N.G. & Allainé, D. 2007. Extra-pair paternity in alpine marmots, *Marmota marmota*: genetic quality and genetic diversity effects. *Behav. Ecol. Sociobiol.* **61**: 1081–1092.

Cohas, A., Yoccoz, N.G., Bonenfant, C., Goossens, B., Genton, C., Galan, M. *et al.* 2008. The genetic similarity between pair members influences the frequency of extrapair paternity in alpine marmots. *Anim. Behav.* **76**: 87–95.

Crew, M.D., Filipowsky, M.E., Zeller, E.C., Smith, G.S. & Walford, R.L. 1990. Major histocompatibility complex class I genes of *Peromyscus leucopus*. *Immunogenetics* **32**: 371–379.

Doherty, P.C. & Zinkernagel, R.M. 1975. Enhanced immunological surveillance in mice heterozygous at the H-2 gene complex. *Nature* **256**: 50–52.

Excoffier, L. & Lischer, H. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **10**: 564–567.

Figueroa, F., Gutknecht, J., Tichy, H. & Klein, J. 1990. Class II MHC genes in rodent evolution. *Immunol. Rev.* **113**: 27–46.

Froeschke, G. & Sommer, S. 2005. MHC class II DRB variability and parasite load in the striped mouse (*Rhabdomys pumilio*) in the Southern Kalahari. *Mol. Biol. Evol.* **22**: 1254–1259.

Garrigan, D. & Hedrick, P.W. 2003. Perspective: detecting adaptive molecular polymorphism: lessons from the MHC. *Evolution* **57**: 1707–1722.

Goossens, B., Chikhi, L., Taberlet, P., Waits, L.P. & Allainé, D. 2001. Microsatellite analysis of genetic variation among and within Alpine marmot populations in the French Alps. *Mol. Ecol.* **10**: 41–52.

Goudet, J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3) [www document]. URL http://www2.unil.ch/popgen/softwares/fstat.htm.

Goüy de Bellocq, J., Charbonnel, N. & Morand, S. 2008. Coevolutionary relationship between helminth diversity and MHC class II polymorphism in rodents. *J. Evol. Biol.* **21**: 1144–1150.

Günther, E. & Walter, L. 2000. Comparative genomic aspects of rat, mouse and human MHC class I gene regions. *Cytogenet. Cell Genet.* **91**: 107–112.

Hedrick, P.W. 1992. Female choice and variation in the major histocompatibility complex. *Genetics* **132**: 575–581.

Hedrick, P.W. 2002. Pathogen resistance and genetic variation at MHC loci. *Evolution* **56**: 1902–1908.

Hughes, A.L. & Nei, M. 1988. Pattern of nucleotide substitution at major histocompatibility complex loci reveals overdominant selection. *Nature* **335**: 167–170.

Hughes, A.L. & Nei, M. 1989. Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. *Proc. Natl Acad. Sci. USA* **86**: 958–962.

Iwasaki, A. & Medzhitov, R. 2010. Regulation of adaptive immunity by the innate immune system. *Science* **327**: 291–295.

Janeway, C.A., Travers, P., Walport, D. & Shlomchik, M.J. 2004. *Immunobiology: The Immune System in Health and Disease*. Garland Publishing, New York, NY.

Kelley, J., Walter, L. & Trowsdale, J. 2005. Comparative genomics of major histocompatibility complexes. *Immunogenetics* **56**: 683–695.

Khazand, M., Peiberg, C., Nagy, M. & Sauermann, U. 1999. *Mhc-DQ-DRB* haplotype analysis in the rhesus macaque: evidence for a number of different haplotypes displaying a low allelic polymorphism. *Tissue Antigens* **54**: 615–624.

Klein, J., Bontrop, R.E., Dawkins, R.L., Erlich, H.A., Gyllensten, U.B., Heise, E.R. *et al.* 1990. Nomenclature for the major histocompatibility complexes of different species: a proposal. *Immunogenetics* **31**: 217–219.

Klein, J., Sato, A., Nagl, S. & O'hUigin, C. 1998. Molecular trans-species polymorphism. *Annu. Rev. Ecol. Syst.* **29**: 1–21.

Kruiswijk, C.P., Hermsen, T., van Heerwaarden, J., Dixon, B., Savelkoul, H.F.J. & Stet, R.J.M. 2005. Major histocompatibility genes in the Lake Tana African large barb species flock: evidence for complete partitioning of class IIB, but not class I, genes among different species. *Immunogenetics* **56**: 894–908.

Kumánovics, A., Madan, A., Qin, S., Rowen, L., Hood, L. & Fischer Lindahl, K. 2002. QUOD ERAT FACIENDUM: sequence analysis of the H2-D and H2-Q regions of 129/SvJ mice. *Immunogenetics* **54**: 479–489.

Leinders-Zufall, T., Brennan, P., Widmayer, P., Chandramani, P., Maul-Pavicic, A., Jäger, M. *et al.* 2004. MHC class I peptides as chemosensory signals in the vomeronasal organ. *Science* **306**: 1033–1037.

Lobigs, M., Rothenfluh, H.S., Blanden, R.V. & Müllbacher, A. 1995. Polymorphic peptide transporters in MHC class I monomorphic Syrian hamster. *Immunogenetics* **42**: 398–407.

Nei, M. & Gojobori, T. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* **3**: 418–426.

Nei, M. & Kumar, S. 2000. *Molecular Evolution and Phylogenetics*. Oxford University Press, New York, NY.

Nizetić, D., Figueroa, F., Dembić, Z., Nevo, E. & Klein, J. 1987. Major histocompatibility complex gene organization in the mole rat *Spalax ehrenbergi*: evidence for transfer of function between class II genes. *Proc. Natl Acad. Sci. USA* **84**: 5828–5832.

Penn, D. & Potts, W. 1998. MHC-disassortative mating preferences reversed by cross-fostering. *Proc. R. Soc. Lond. B* **265**: 1299–1306.

Piertney, S.B. & Oliver, M.K. 2006. The evolutionary ecology of the major histocompatibility complex. *Heredity* **96**: 7–21.

Posada, D. & Buckley, T.R. 2004. Model selection and model averaging in phylogenetics: advantages of Akaike information

criterion and Bayesian approaches over likelihood ratio tests. *Syst. Biol.* **53**: 793–808.
Preleuthner, M. & Pinsker, W. 1993. Depauperated gene pools in *Marmota m. marmota* are caused by an ancient bottleneck: electrophoretic analysis of wild populations from Austria and Switzerland. *Acta Theriol.* **38**: 121–139.
Promerova, M., Albrecht, T. & Bryja, J. 2009. Extremely high MHC class I variation in a population of a long-distance migrant, the Scarlet Rosefinch (*Carpodacus erythrinus*). *Immunogenetics* **61**: 451–461.
Radwan, J., Biedrzycka, A. & Babik, W. 2010. Does reduced MHC diversity decrease viability of vertebrate populations? *Biol. Conserv.* **143**: 537–544.
Rassmann, K., Arnold, W. & Tautz, D. 1994. Low genetic variability in a natural alpine marmot population (*Marmota marmota*, Sciuridae) revealed by DNA fingerprinting. *Mol. Ecol.* **3**: 347–353.
Reche, P.A. & Reinherz, E.L. 2003. Sequence variability analysis of human class I and class II MHC molecules: functional and structural correlates of amino acid polymorphisms. *J. Mol. Biol.* **331**: 623–641.
Ricanova, S., Bryja, J., Cosson, J.-F., Csongor, G., Choleva, L., Ambros, M. *et al.* 2011. Depleted genetic variation of the European ground squirrel in Central Europe in both microsatellites and the major histocompatibility complex gene: implications for conservation. *Conserv. Genet.*, **12**: 1115–1129.
Richman, A.D., Herrera, L.G. & Nash, D. 2001. MHC class II beta sequence diversity in the deer mouse (*Peromyscus maniculatus*): implications for models of balancing selection. *Mol. Ecol.* **10**: 2765–2773.
Ronquist, F. & Huelsenbeck, J.P. 2003. MrBayes 3: bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**: 1572–1574.
Shum, B., Guethlein, L., Flodin, L., Adkison, M., Hedrick, R., Nehring, R. *et al.* 2001. Modes of salmonid MHC class I and II evolution differ from the primate paradigm. *J. Immunol* **166**: 3297–3308.
Snell, G.D. 1968. The H-2 locus of the mouse: observations and speculations concerning its comparative genetics and its polymorphism. *Folia Biol. (Prague)* **14**: 335–358.
Spurgin, L.G. & Richardson, D.S. 2010. How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proc. R. Soc. Lond. B* **277**: 979–988.
Takahata, N. & Nei, M. 1990. Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. *Genetics* **124**: 967–978.
Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. & Kumar, S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**: 2731–2739.
Tollenaere, C., Bryja, J., Galan, M., Cadet, P., Deter, J., Chaval, Y. *et al.* 2008. Multiple parasites mediate balancing selection at two MHC class II genes in the fossorial water vole: insights from multivariate analyses and population genetics. *J. Evol. Biol.* **21**: 1307–1320.
Wegner, K.M. 2008. Historical and contemporary selection of teleost MHC genes: did we leave the past behind? *J. Fish Biol.* **73**: 2110–2132.
Wettstein, P.J., Strausbauch, M., Johnston, S.L. & States, J. 1996. The divergence of major histocompatibility complex class I genes in *Sciurus aberti*. *Mol. Biol. Evol.* **13**: 56–66.
Yang, Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**: 1586–1591.
Yang, D.L., Lu, M., Hao, L.J. & Roggendorf, M. 2000. Molecular cloning and characterization of major histocompatibility complex class I cDNAs from woodchuck (*Marmota monax*). *Tissue Antigens* **55**: 548–557.
Yeager, M. & Hughes, A.L. 1999. Evolution of the mammalian MHC: natural selection, recombination, and convergent evolution. *Immunol. Rev.* **167**: 45–58.
Zagalska-Neubauer, M., Babik, W., Stuglik, M., Gustafsson, L., Cichoń, M. & Radwan, J. 2010. 454 sequencing reveals extreme complexity of the class II Major Histocompatibility Complex in the collared flycatcher. *BMC Evol. Biol.* **10**: 395.
Zhang, J., Nielsen, R. & Yang, Z. 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol. Biol. Evol.* **22**: 2472–2479.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Table S1 Primer sequences for the PCR amplification of MHC class I and class II DRB loci in the Alpine marmot (*Marmota marmota*).

Table S2 The PCR conditions and polymorphism characterization of MHC class I and class II DRB loci in *Marmota marmota*.

Table S3 The allele frequencies of four polymorphic loci in *Marmota marmota*.

Table S4 Observed (Ho), expected (He) heterozygosities and allelic richness (R) per putative locus per population.

Table S5 The amino acid sequences of MHC class I second exon alleles (a) and of MHC class II DRB second exon alleles (b) in *Marmota marmota*.

Table S6 The average pairwise nucleotide distances (Kimura 2-parameter models, K2P), Poisson-corrected amino acid distances, the average rates of synonymous (d_S) and nonsynonymous (d_N) substitutions per site and the results of *Z*-test of positive selection for polymorphic putative MHC loci in the Alpine marmot.

Table S7 The evaluation of the fit of various models of codon evolution.

Table S8 *Marmota marmota* MHC class I sequences.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be reorganized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

Received 1 February 2012; revised 6 April 2012; accepted 12 April 2012